

CLAIMS:

I claim:

1. A self-tuning cache comprising:

a primary cache;

at least two test caches, a first one of said test caches having a cache size which is smaller than a size of said primary cache, and a second one of said test caches having a cache size which is greater than said size of said primary cache;

a cache engine programmed to manage said primary cache and said at least two test caches; and,

a cache tuner coupled to said primary and test caches, said cache tuner comprising a configuration for resizing said primary cache when one of said at least two test caches demonstrates cache performance which justifies resizing said primary cache.

2. The self-tuning cache of claim 1, wherein each of said at least two test caches comprise a configuration for storing cache keys for cacheable objects and corresponding placeholders for said cacheable objects in lieu of storing said cacheable objects.

3. The self-tuning cache of claim 1, wherein said first one of said test caches comprises a cache size which is half that of said primary cache.

4. The self-tuning cache of claim 3, wherein said second one of said test caches comprises a cache size which is double that of said primary cache.
5. The self-tuning cache of claim 1, further comprising a maximum limit and a minimum limit for resizing said primary cache.
6. A method for self-tuning an active cache, the method comprising the steps of:
managing the active cache by inserting, retrieving and evicting cacheable objects and corresponding caching keys in the active cache and by locating cached objects selected for retrieval from the active cache by reference to corresponding ones of said caching keys;
further managing a test cache by inserting and evicting in said test cache caching keys and dummy placeholders for cacheable objects not stored in said test cache and by locating in said test cache individual ones of said caching keys corresponding to requested ones of said cacheable objects;
measuring and comparing hit rates for each of said active cache and said test cache; and,
if said measured hit rates compare such that a change in size for the active cache is justified, resizing the active cache and said test cache.
7. The method of claim 6, wherein said resizing step comprises:

if said test cache is smaller in size than the active cache and if said test cache demonstrates a hit rate which does not differ significantly from a hit rate measured for the active cache, resizing the active cache to a smaller size.

8. The method of claim 6, wherein said resizing step comprises:

if said test cache is larger in size than the active cache and if said test cache demonstrates a hit rate which significantly exceeds a hit rate measured for the active cache, resizing the active cache to a larger size.

9. The method of claim 6, further comprising the step of limiting said resizing so as to not exceed a minimum and a maximum cache size for the active cache.

10. The method of claim 6, further comprising the step of rearranging a data structure for the active cache based upon a change in size for the active cache.

11. A method for self-tuning an active cache, the method comprising the steps of:

receiving a request to retrieve an object;

generating a cache key for said object;

searching the active cache for said object using said generated cache key;

further searching at least one test cache for a stored cache key which matches said generated cache key;

returning said object from the active cache if said object is located in the active cache in said searching step;

updating hit rate statistics for each of the active cache and said at least one test cache based upon whether said object is located in the active cache in said searching step, and whether said generated cache key matches a stored cache key in said at least one test cache; and,

determining whether to resize the active cache based upon said updated hit rate statistics.

12. The method of claim 11, wherein said determining step comprises the step of:

if said at least one test cache is smaller in size than the active cache and if said at least one test cache demonstrates a hit rate which does not differ significantly from a hit rate measured for the active cache, resizing the active cache to a smaller size.

13. The method of claim 11, wherein said determining step comprises the step of:

if said at least one test cache is larger in size than the active cache and if said at least one test cache demonstrates a hit rate which significantly exceeds a hit rate measured for the active cache, resizing the active cache to a larger size.

14. The method of claim 11, further comprising the step of evicting stored cache keys from said at least one test cache.

15. The method of claim 11, further comprising the step of inserting a generated cache key into said at least one test cache.

16. A machine readable storage having stored thereon a computer program for self-tuning an active cache, the computer program comprising a routine set of instructions which when executed by a machine cause the machine to perform the steps of:

managing the active cache by inserting, retrieving and evicting cacheable objects and corresponding caching keys in the active cache and by locating cached objects selected for retrieval from the active cache by reference to said caching keys;

further managing a test cache by inserting and evicting in said test cache caching keys and dummy placeholders for cacheable objects not stored in said test cache and by locating in said test cache individual ones of said caching keys corresponding to requested ones of said cacheable objects;

measuring and comparing hit rates for each of said active cache and said test cache; and,

if said measured hit rates compare such that a change in size for the active cache is justified, resizing the active cache and said test cache.

17. The machine readable storage of claim 16, wherein said resizing step comprises:

if said test cache is smaller in size than the active cache and if said test cache demonstrates a hit rate which does not differ significantly from a hit rate measured for the active cache, resizing the active cache to a smaller size.

18. The machine readable storage of claim 16, wherein said resizing step comprises:

if said test cache is larger in size than the active cache and if said test cache demonstrates a hit rate which significantly exceeds a hit rate measured for the active cache, resizing the active cache to a larger size.

19. The machine readable storage of claim 16, further comprising the step of limiting said resizing so as to not exceed a minimum and a maximum cache size for the active cache.

20. The machine readable storage of claim 16, further comprising the step of rearranging a data structure for the active cache based upon a change in size for the active cache.

21. A machine readable storage having stored thereon a computer program for self-tuning an active cache, the computer program comprising a routine set of instructions which when executed by a machine cause the machine to perform the steps of:

- receiving a request to retrieve an object;
- generating a cache key for said object;
- searching the active cache for said object using said generated cache key;
- further searching at least one test cache for a stored cache key which matches said generated cache key;
- returning said object from the active cache if said object is located in the active cache in said searching step;

updating hit rate statistics for each of the active cache and said at least one test cache based upon whether said object is located in the active cache in said searching step, and whether said generated cache key matches a stored cache key in said at least one test cache; and,

determining whether to resize the active cache based upon said hit rate statistics.

22. The machine readable storage of claim 21, wherein said determining step comprises the step of:

if said at least one test cache is smaller in size than the active cache and if said at least one test cache demonstrates a hit rate which does not differ significantly from a hit rate measured for the active cache, resizing the active cache to a smaller size.

23. The machine readable storage of claim 21, wherein said determining step comprises the step of:

if said at least one test cache is larger in size than the active cache and if said at least one test cache demonstrates a hit rate which significantly exceeds a hit rate measured for the active cache, resizing the active cache to a larger size.

24. The machine readable storage of claim 21, further comprising the step of evicting stored cache keys from said at least one test cache.

25. The machine readable storage of claim 16, further comprising the step of inserting a generated cache key into said at least one test cache.